

WO 99/38967

PCT/EP99/00478

1
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: DOMPE' S.p.A.
(B) STREET: Via Campo di Pile
(C) CITY: L'AQUILA
(E) COUNTRY: ITALY
(F) POSTAL CODE (ZIP): 67100

(ii) TITLE OF INVENTION: RECOMBINANT PROTEINS DERIVED FROM HGF AND MSP

(iii) NUMBER OF SEQUENCES: 4

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1725 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGTGGGTGA CCAAACTCCT GCCAGCCCTG CTGCTGCAGC ATGTCCTCCT GCATCTCCTC 60
CTGCTCCCCA TCGCCATCCC CTATGCAGAG GGACAAAGGA AAAGAAGAAA TACAATTTCAT 120
GAATTCAAAA AATCAGCAAA GACTACCCTA ATCAAAATAG ATCCAGCACT GAAGATAAAA 180
ACCAAAAAAG TGAATACTGC AGACCAATGT GCTAATAGAT GTACTAGGAA TAAAGGACTT 240
CCATTCACCTT GCAAGGCTTT TGTTTTTGAT AAAGCAAGAA AACAATGCCT CTGGTTCCCC 300
TTCAATAGCA TGTCAAGTGG AGTGAAAAAA GAATTTGGCC ATGAATTTGA CCTCTATGAA 360
AACAAAGACT ACATTAGAAA CTGCATCATT GGTAAGGAC GCAGCTACAA GGGAACAGTA 420
TCTATCACTA AGAGTGGCAT CAAATGTCAG CCCTGGAGTT CCATGATACC ACACGAACAC 480
AGCTATCGGG GTAAAGACCT ACAGGAAAAC TACTGTGCGAA ATCCTCGAGG GGAAGAAGGG 540
GGACCCTGGT GTTTCACAAG CAATCCAGAG GTACGCTACG AACTCTGTGA CATTCTCAG 600

pub
B11
Cont. 1

005550" 03300

TGTTTCAGAAG TTGAATGCAT GACCTGCAAT GGGGAGAGTT ATCGAGGTCT CATGGATCAT 660
ACAGAATCAG GCAAGATTTG TCAGCGCTGG GATCATCAGA CACCACACCG GCACAAATTC 720
TTGCCTGAAA GATATCCCGA CAAGGGCTTT GATGATAATT ATTGCCGCAA TCCCGATGGC 780
CAGCCGAGGC CATGGTGCTA TACTCTTGAC CCTCACACCC GCTGGGAGTA CTGTGCAATT 840
AAAACATGCG CTGACAAAGC TTCGGGCGGT GCGGTTCTG GTGGCGGTGG CTCCGGCGGT 900
GGCGGTTCTC TAGAGGGACA AAGGAAAAGA AGAAATACAA TTCATGAATT CAAAAAATCA 960
GCAAAGACTA CCCTAATCAA AATAGATCCA GCACTGAAGA TAAAAACCAA AAAAGTGAAT 1020
ACTGCAGACC AATGTGCTAA TAGATGTACT AGGAATAAAG GACTTCCATT CACTTGCAAG 1080
GCTTTTGTTC TTGATAAAGC AAGAAAACAA TGCCTCTGGT TCCCCTTCAA TAGCATGTCA 1140
AGTGGAGTGA AAAAAGAATT TGGCCATGAA TTTGACCTCT ATGAAAACAA AGACTACATT 1200
AGAAACTGCA TCATTGGTAA AGGACGCAGC TACAAGGGAA CAGTATCTAT CACTAAGAGT 1260
GGCATCAAAT GTCAGCCCTG GAGTTCCATG ATACCACACG AACACAGCTA TCGGGGTAAA 1320
GACCTACAGG AAAACTACTG TCGAAATCCT CGAGGGGAAG AAGGGGGACC CTGGTGTTTC 1380
ACAAGCAATC CAGAGGTACG CTACGAAGTC TGTGACATTC CTCAGTGTTT AGAAGTTGAA 1440
TGCATGACCT GCAATGGGGA GAGTTATCGA GGTCTCATGG ATCATAACAG ATCAGGCAAG 1500
ATTTGTACAG GCTGGGATCA TCAGACACCA CACCGGCACA AATTCTTGCC TGAAAGATAT 1560
CCCGACAAGG GCTTTGATGA TAATTATTGC CGCAATCCCG ATGGCCAGCC GAGGCCATGG 1620
TGCTATACTC TTGACCCTCA CACCCGCTGG GAGTACTGTG CAATTAAAC ATGCGCTGAC 1680
AAAGCTGACG ACGACGACAA ACACCACCAC CACCACCACC ACTAG 1725

3

Sub
Bill
cont'd

(A) LENGTH: 574 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

Секрет

Cys Asn Gly Glu Ser Tyr Arg Gly Leu Met Asp His Thr Glu Ser Gly
210 215 220

5

Tyr Cys Arg Asn Pro Asp Gly Gln Pro Arg Pro Trp Cys Tyr Thr Leu
530 535 540

Asp Pro His Thr Arg Trp Glu Tyr Cys Ala Ile Lys Thr Cys Ala Asp
545 550 555 560

Lys Ala Asp Asp Asp Asp Lys His His His His His His His
565 570

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1692 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATGGGGTGGC TCCCACTCCT GCTGCTTCTG ACTCAATGCT TAGGGGTCCC TGGGCAGCGC 60
TCGCCATTGA ATGACTTCCA AGTGCTCCGG GGCACAGAGC TACAGCACCT GCTACATGCG 120
GTGGTGCCCG GGCCTTGCCA GGAGGATGTG GCAGATGCTG AAGAGTGTGC TGGTCGCTGT 180
GGGCCCTTAA TGGACTGCCG GGCCTTCCAC TACAACGTGA GCAGCCATGG TTGCCAACTG 240
CTGCCATGGA CTCAACACTC GCCCCACACG AGGCTGCGGC GTTCTGGGCG CTGTGACCTC 300
TTCCAGAAGA AAGACTACGT ACGGACCTGC ATCATGAACA ATGGGGTTGG GTACCGGGGC 360
ACCATGGCCA CGACCGTGGG TGGCCTGCCC TGCCAGGCTT GGAGCCACAA GTTCCCGAAT 420
GATCACAAGT ACACGCCCAC TCTCCGGAAT GGCCTGGAAG AGAACTTCTG CCGTAACCCT 480
GATGGCGACC CCGGAGGTCC TTGGTGCTAC ACAACAGACC CTGCTGTGCG CTTCCAGAGC 540
TGCGGCATCA AATCCTGCCG GGAGGCCGCG TGTGTCTGGT GCAATGGCGA GGAATACCGC 600
GGCGCGGTAG ACCGCACGGA GTCAGGGCGC GAGTGCCAGC GCTGGGATCT TCAGCACCCG 660
CACCAGCACC CCTTCGAGCC GGGCAAGTTC CTCGACCAAG GTCTGGACGA CAACTATTGC 720
CGGAATCCTG ACGGCTCCGA GCGGCCATGG TGCTACACTA CGGATCCGCA GATCGAGCGA 780
GAGTTCTGTG ACCTCCCCCG CTGCGGGTCC GAGGCACAGC CCCGCCTCGA GGGCGGTGGC 840
GGTTCTGGTG GCGGTGGCTC CGGCGGTGGC GGTTCCTTAG AGGGACAAAG GAAAAGAAGA 900
AATACAATTC ATGAATTCAA AAAATCAGCA AAGACTACCC TAATCAAAAT AGATCCAGCA 960

Pub
B11
Cont'd

(2) INFORMATION FOR SEQ ID NO: 4:

(A) LENGTH: 563 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Asp Cys Arg Ala Phe His Tyr Asn Val Ser Ser His Gly Cys Gln Leu
65 70 75 80

Leu	Pro	Trp	Thr	Gln	His	Ser	Pro	His	Thr	Arg	Leu	Arg	Arg	Ser	Gly
				85					90					95	
Arg	Cys	Asp	Leu	Phe	Gln	Lys	Lys	Asp	Tyr	Val	Arg	Thr	Cys	Ile	Met
			100					105					110		
Asn	Asn	Gly	Val	Gly	Tyr	Arg	Gly	Thr	Met	Ala	Thr	Thr	Val	Gly	Gly
		115					120					125			
Leu	Pro	Cys	Gln	Ala	Trp	Ser	His	Lys	Phe	Pro	Asn	Asp	His	Lys	Tyr
	130					135					140				
Thr	Pro	Thr	Leu	Arg	Asn	Gly	Leu	Glu	Glu	Asn	Phe	Cys	Arg	Asn	Pro
145					150					155					160
Asp	Gly	Asp	Pro	Gly	Gly	Pro	Trp	Cys	Tyr	Thr	Thr	Asp	Pro	Ala	Val
				165					170					175	
Arg	Phe	Gln	Ser	Cys	Gly	Ile	Lys	Ser	Cys	Arg	Glu	Ala	Ala	Cys	Val
			180					185					190		
Trp	Cys	Asn	Gly	Glu	Glu	Tyr	Arg	Gly	Ala	Val	Asp	Arg	Thr	Glu	Ser
		195					200					205			
Gly	Arg	Glu	Cys	Gln	Arg	Trp	Asp	Leu	Gln	His	Pro	His	Gln	His	Pro
	210					215					220				
Phe	Glu	Pro	Gly	Lys	Phe	Leu	Asp	Gln	Gly	Leu	Asp	Asp	Asn	Tyr	Cys
225					230					235					240
Arg	Asn	Pro	Asp	Gly	Ser	Glu	Arg	Pro	Trp	Cys	Tyr	Thr	Thr	Asp	Pro
				245					250					255	
Gln	Ile	Glu	Arg	Glu	Phe	Cys	Asp	Leu	Pro	Arg	Cys	Gly	Ser	Glu	Ala
			260					265					270		
Gln	Pro	Arg	Leu	Glu	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly
		275					280					285			
Gly	Gly	Gly	Ser	Leu	Glu	Gly	Gln	Arg	Lys	Arg	Arg	Asn	Thr	Ile	His
	290					295					300				
Glu	Phe	Lys	Lys	Ser	Ala	Lys	Thr	Thr	Leu	Ile	Lys	Ile	Asp	Pro	Ala
305					310					315					320
Leu	Lys	Ile	Lys	Thr	Lys	Lys	Val	Asn	Thr	Ala	Asp	Gln	Cys	Ala	Asn
				325					330					335	
Arg	Cys	Thr	Arg	Asn	Lys	Gly	Leu	Pro	Phe	Thr	Cys	Lys	Ala	Phe	Val
			340					345					350		
Phe	Asp	Lys	Ala	Arg	Lys	Gln	Cys	Leu	Trp	Phe	Pro	Phe	Asn	Ser	Met
		355					360					365			
Ser	Ser	Gly	Val	Lys	Lys	Glu	Phe	Gly	His	Glu	Phe	Asp	Leu	Tyr	Glu
	370					375					380				

sub
B1
Cont'd

THE UNIVERSITY OF CHICAGO

